

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Peter DORMER
- (ii) TITLE OF INVENTION: PROTEIN WITH DIFFERENTIATION-INDUCING  
ACTIVITY FOR FRIEND'S ERYTHROLEUKEMIA CELL LINES
- (iii) NUMBER OF SEQUENCES: 10
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER
  - (B) STREET: 99 Canal Center Plaza, Suite 300
  - (C) CITY: Alexandria
  - (D) STATE: VA
  - (E) COUNTRY: USA
  - (F) ZIP: 22314
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Presta, Frank P.
  - (B) REGISTRATION NUMBER: 19,828
  - (C) REFERENCE/DOCKET NUMBER: 3428-005
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (703) 684-1111
  - (B) TELEFAX: (703) 684-1124

(2) INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1495 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Mus musculus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CCGACCGTGC GGA	CTTAAGA TGGAGGCACT TCCTGTCTGC GGC	GGGAAGA GAAGGCTCGG	60
TCGGAGCCCGG GAATGCTGGG ACTTG	TAGTG CAGTAGTCAAT GGT	TCTCTAT GGGCTTTCAG	120

AGTGAGTGGC GGGGAAGGCGG CCCCAGAGGCA TGCTGGGAGT TGTAGTCCTG CCGTCGTCAA	180
TGGTTCTCTA TGGGCTTTCA GAGTGAGTGG CGGGAAGGCG GCCCCGAGGC ATGCTGGGAG	240
TTGTAGTCCT GCCATAGTCA ATGGTTCTCT ATGGGCTTTC AGACTGAGTG GCGGGAAGGC	300
GGCCCCGAGG CATGCTGGGA GTTGCAGCGC CATGTTTTAA AGCACGCGTT TCTCTGTATA	360
GACCTGGCTG TGGATTTTTT GCTAATTCTT TTTTTAGCT TTATTTTTAA TTTTACTTT	420
TTACACAGG ATTTCTCTTT ATAGCCTTGG CTACCGTTTT TTCCCTAATT ATTCTCCTTT	480
TCATTTTGGT TTATTTTTTT TTAATTTTGG TTTTTTAA ACAGGGTTTC TCTGTATAGA	540
CCTGGCTGTG GATTTCTCAC TAATTATTTT TTTTAGCTTT ATTTTAAATT TTTACTTTTT	600
CACACAGGAT TTCTCTTTAT AGCCTTGGCT ACCGTTTTTT CCGTAATTAT TCTTATTTTC	660
ATTTTGGTTT ATTTTTTAAT TTTAATTTT GATTTTGGAG ACAGGGTTTC TCTTTTAGCC	720
GCAGCTATGG TTTCTGCCCT AATTATTCTT GTCCTTATTT GTAATTTAAT TCTTAATTTA	780
ATTTAATTTA TAATTTTGTT GTAAGTTTTT CTGTGGGCGT GAATGGAAAG TCTAACCCGT	840
GTTTCTCTGT TCAGCGTCCG CCGGTCACGG CCGCCGCCCC CAGCGACGTC ACCCACACGC	900
GCAGAAGCGG ACGCCGCGGT CAAGATGTCT CTGCCATGCC CACGGGACGC ACGGACGCAC	960
GGACGGACGG ACGGACTCCA CAAGGTAGGA AGCCTGCGCC GACCGCACCG CCGCACCCAC	1020
CACAGCACAC AGGACACACG CGGGCCCCGC GCCCGCCCAG GCACACGCGG CACACACGGC	1080
ACACACGGCA GGCAGGCCAG GCACACGCAT CCGCAGGACC CGCCGCACCC GCCACGCAGA	1140
CACGGACGAG CCGCCGCGGT CAAGATGTTT ACCCGCCGCG GTCAAGATGT ATGTGCCACC	1200
GACCCTCGCC CCGCTGGACG GACGGACGGA CGCACGCACG CCGTCAGCGT CCACCGGTCA	1260
CTGCCGCCGC CCACAGTGAT GTCACCCACG AAAGCACACA CGTAGAAGCG GACGCCGTGG	1320
TCAAGATGTC TCTGCCATCC CCACAGGACG GACGGACGGA CTCCACAAGG TGGCGGTGTC	1380
GCCGAGGCCG CCAGGACGGA GCGATTCTCA CGGAGGAAGG AGCACGCCAA CAGGGCCTGA	1440
CTGCGTACAG ACATGTCCCC CTCAATAAAA TTGCAGTTGA AATGGAAAAA AAAAA	1495

(2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 715 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: YES

- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 155..688

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

CGCGCCCGCC CGGGATCCCC AGCTGCCGCC GCGCCCGCCC GCCCGCCCGG GGCCCCCGCT	60
---	----

GCAGAACCGT GACCGTCCGC CGGTCACGGC CGCCGCCCC AGCGACGTCA CCCACACGG	120
CAGAAGCGGA CGCCGCGGTC AAGATGTCTC TGCC ATG CCC ACG GGA CGC ACG	172
Met Pro Thr Gly Arg Thr	
1 5	
GAC GCA CGG ACG GAC GGA CTG ACT CCA CAA GGT AGG AAG CCT GCG CCG	220
Asp Ala Arg Thr Asp Gly Leu Thr Pro Gln Gly Arg Lys Pro Ala Pro	
10 15 20	
ACC GCA CCG CCG CAC CCA CCA CAG CAC ACA GGA CAC ACG CGG GCC CCG	268
Thr Ala Pro Pro His Pro Pro Gln His Thr Gly His Thr Arg Ala Pro	
25 30 35	
CGC CCG CCC AGG CAC ACG CGG CAC ACA CGG CAC ACA CGG CAG GCA GGC	316
Arg Pro Pro Arg His Thr Arg His Thr Arg His Thr Arg Gln Ala Gly	
40 45 50	
CAG GCA CAC GCA TCC GCA GGA CCC GCC GCA CCC GCC ACG CAG ACA CGG	364
Gln Ala His Ala Ser Ala Gly Pro Ala Ala Pro Ala Thr Gln Thr Arg	
55 60 65 70	
ACG AGC CGC CGC GGT CAA GAT GTT CAC CCG CCG CGG TCA AGA TGT ATG	412
Thr Ser Arg Arg Gly Gln Asp Val His Pro Pro Arg Ser Arg Cys Met	
75 80 85	
TGC CAC CGA CCC TCG CCC CGC TGG ACG GAC GGA CGG ACG CGC GCA CGC	460
Cys His Arg Pro Ser Pro Arg Trp Thr Asp Gly Arg Thr Arg Ala Arg	
90 95 100	
CGT CAG CGT CCA CCG GTC ACT GCC GCC GCC CAC AGT GAC GTC ACC CAC	508
Arg Gln Arg Pro Pro Val Thr Ala Ala Ala His Ser Asp Val Thr His	
105 110 115	
GAA AGC ACA CAC GTA GAA GCG GAC GCC GTG GTC AAG ATG TCT CTG CCA	556
Glu Ser Thr His Val Glu Ala Asp Ala Val Val Lys Met Ser Leu Pro	
120 125 130	
TCC CCA CAG GAC GGA CGG ACG GAC TCC ACA AGG TGC GCG TGT CGC CGA	604
Ser Pro Gln Asp Gly Arg Thr Asp Ser Thr Arg Cys Ala Cys Arg Arg	
135 140 145 150	
GGC CGC CAG GAT GGA GCG ATT CTC ACG GAG GAA GGA GCA CGC CAA CAG	652
Gly Arg Gln Asp Gly Ala Ile Leu Thr Glu Glu Gly Ala Arg Gln Gln	
155 160 165	
GGC CTG ACT GCG TAC AGA AAT GCC CCC CCT CAA TAA AATTGCAGTT	698
Gly Leu Thr Ala Tyr Arg Asn Ala Pro Pro Gln *	
170 175	
GAAATGGAAA AAAAAAA	715

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 177 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Met Pro Thr Gly Arg Thr Asp Ala Arg Thr Asp Gly Leu Thr Pro Gln
1 5 10 15
Gly Arg Lys Pro Ala Pro Thr Ala Pro Pro His Pro Pro Gln His Thr

	20		25		30										
Gly	His	Thr	Arg	Ala	Pro	Arg	Pro	Pro	Arg	His	Thr	Arg	His	Thr	Arg
	35						40					45			
His	Thr	Arg	Gln	Ala	Gly	Gln	Ala	His	Ala	Ser	Ala	Gly	Pro	Ala	Ala
	50					55					60				
Pro	Ala	Thr	Gln	Thr	Arg	Thr	Ser	Arg	Arg	Gly	Gln	Asp	Val	His	Pro
	65				70					75					80
Pro	Arg	Ser	Arg	Cys	Met	Cys	His	Arg	Pro	Ser	Pro	Arg	Trp	Thr	Asp
				85					90					95	
Gly	Arg	Thr	Arg	Ala	Arg	Arg	Gln	Arg	Pro	Pro	Val	Thr	Ala	Ala	Ala
			100					105					110		
His	Ser	Asp	Val	Thr	His	Glu	Ser	Thr	His	Val	Glu	Ala	Asp	Ala	Val
		115					120					125			
Val	Lys	Met	Ser	Leu	Pro	Ser	Pro	Gln	Asp	Gly	Arg	Thr	Asp	Ser	Thr
	130					135					140				
Arg	Cys	Ala	Cys	Arg	Arg	Gly	Arg	Gln	Asp	Gly	Ala	Ile	Leu	Thr	Glu
	145				150					155					160
Glu	Gly	Ala	Arg	Gln	Gln	Gly	Leu	Thr	Ala	Tyr	Arg	Asn	Ala	Pro	Pro
				165				170						175	

Gln

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 636 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: YES

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..636

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

ATG	GGG	CTG	CAG	AAC	CGT	GAC	CGT	CCG	CCG	GTC	ACG	GCC	GCC	GCC	CCC	48
Met	Gly	Leu	Gln	Asn	Arg	Asp	Arg	Pro	Pro	Val	Thr	Ala	Ala	Ala	Pro	
	180					185					190					
AGC	GAC	GTC	ACC	CAC	ACG	CGC	AGA	AGC	GGA	CGC	CGC	GGT	CAA	GAT	GTC	96
Ser	Asp	Val	Thr	His	Thr	Arg	Arg	Ser	Gly	Arg	Arg	Gly	Gln	Asp	Val	
	195				200				205						210	
TCT	GCC	ATG	CCC	ACG	GGA	CGC	ACG	GAC	GCA	CGG	ACG	GAC	GGA	CTG	ACT	144
Ser	Ala	Met	Pro	Thr	Gly	Arg	Thr	Asp	Ala	Arg	Thr	Asp	Gly	Leu	Thr	
				215				220						225		
CCA	CAA	GGT	AGG	AAG	CCT	GCG	CCG	ACC	GCA	CCG	CCG	CAC	CCA	CCA	CAG	192
Pro	Gln	Gly	Arg	Lys	Pro	Ala	Pro	Thr	Ala	Pro	Pro	His	Pro	Pro	Gln	
				230				235						240		

CAC ACA GGA CAC ACG CGG GCC CCG CGC CCG CCC AGG CAC ACG CGG CAC	240
His Thr Gly His Thr Arg Ala Pro Arg Pro Pro Arg His Thr Arg His	
245 250 255	
ACA CGG CAC ACA CGG CAG GCA GGC CAG GCA CAC GCA TCC GCA GGA CCC	288
Thr Arg His Thr Arg Gln Ala Gly Gln Ala His Ala Ser Ala Gly Pro	
260 265 270	
GCC GCA CCC GCC ACG CAG ACA CGG ACG AGC CGC CGC GGT CAA GAT GTT	336
Ala Ala Pro Ala Thr Gln Thr Arg Thr Ser Arg Arg Gly Gln Asp Val	
275 280 285 290	
CAC CCG CCG CGG TCA AGA TGT ATG TGC CAC CGA CCC TCG CCC CGC TGG	384
His Pro Pro Arg Ser Arg Cys Met Cys His Arg Pro Ser Pro Arg Trp	
295 300 305	
ACG GAC GGA CGG ACG CGC GCA CGC CGT CAG CGT CCA CCG GTC ACT GCC	432
Thr Asp Gly Arg Thr Arg Ala Arg Arg Gln Arg Pro Pro Val Thr Ala	
310 315 320	
GCC GCC CAC AGT GAC GTC ACC CAC GAA AGC ACA CAC GTA GAA GCG GAC	480
Ala Ala His Ser Asp Val Thr His Glu Ser Thr His Val Glu Ala Asp	
325 330 335	
GCC GTG GTC AAG ATG TCT CTG CCA TCC CCA CAG GAC GGA CGG ACG CAC	528
Ala Val Val Lys Met Ser Leu Pro Ser Pro Gln Asp Gly Arg Thr Asp	
340 345 350	
TCC ACA AGG TGC GCG TGT CGC CGA GGC CGC CAG GAT GGA GCG ATT CTC	576
Ser Thr Arg Cys Ala Cys Arg Arg Gly Arg Gln Asp Gly Ala Ile Leu	
355 360 365 370	
ACG GAG GAA GGA GCA CGC CAA CAG GGC CTG ACT GCG TAC AGA AAT GCC	624
Thr Glu Glu Gly Ala Arg Gln Gln Gly Leu Thr Ala Tyr Arg Asn Ala	
375 380 385	
CCC CCT CAA TAA	636
Pro Pro Gln *	
390	

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 211 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Met Gly Leu Gln Asn Arg Asp Arg Pro Pro Val Thr Ala Ala Ala Pro	
1 5 10 15	
Ser Asp Val Thr His Thr Arg Arg Ser Gly Arg Arg Gly Gln Asp Val	
20 25 30	
Ser Ala Met Pro Thr Gly Arg Thr Asp Ala Arg Thr Asp Gly Leu Thr	
35 40 45	
Pro Gln Gly Arg Lys Pro Ala Pro Thr Ala Pro Pro His Pro Pro Gln	
50 55 60	
His Thr Gly His Thr Arg Ala Pro Arg Pro Pro Arg His Thr Arg His	
65 70 75 80	
Thr Arg His Thr Arg Gln Ala Gly Gln Ala His Ala Ser Ala Gly Pro	

85

90

95

Ala Ala Pro Ala Thr Gln Thr Arg Thr Ser Arg Arg Gly Gln Asp Val  
 100 105 110

His Pro Pro Arg Ser Arg Cys Met Cys His Arg Pro Ser Pro Arg Trp  
 115 120 125

Thr Asp Gly Arg Thr Arg Ala Arg Arg Gln Arg Pro Pro Val Thr Ala  
 130 135 140

Ala Ala His Ser Asp Val Thr His Glu Ser Thr His Val Glu Ala Asp  
 145 150 155 160

Ala Val Val Lys Met Ser Leu Pro Ser Pro Gln Asp Gly Arg Thr Asp  
 165 170 175

Ser Thr Arg Cys Ala Cys Arg Arg Gly Arg Gln Asp Gly Ala Ile Leu  
 180 185 190

Thr Glu Glu Gly Ala Arg Gln Gln Gly Leu Thr Ala Tyr Arg Asn Ala  
 195 200 205

Pro Pro Gln  
 210

## (2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 42 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Mus musculus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CGTCCGCCGG TCACGGCCGC CGCCCCCAGC GACGTCACCC AC

42

## (2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 55 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Mus musculus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

AGAAGCGGAC GCCGCGGTCA AGATGTCTCT GCCATGCCCA CGGGACGCAC GGACG

55

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mus musculus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TAGTCCTGCC GTCGTCAATG GTTCTCTATG GGCTTTCAGA GTGAGTGGCG GGAAGCGGGC

60

CCCGAGGCAT GCTGGGAGTT G

81

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 178 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mus musculus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GTTTCTCTGT ATAGACCTGG CTGTGGATTT TTCGCTAATT CTTTTTTTGA GCTTTATTTT

60

TAATTTTAC TTTTTCACAC AGGATTTCCT TTTATAGCCT TGGCTACCGT TTTTCCCTA

120

ATTATTCTCC TTTTCATTTT GGTTTATTTT TTTTAAATTT TGGTTTTTTT AAGACAGG

178

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 385 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mus musculus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

AGGAAGCCTG CGCCGACCGC ACCGCCGCAC CCACCACAGC ACACAGGACA CACGCGGGCC	60
CCGCGCCCCG CCAGGCACAC GCGGCACACA CGGCACACAC GGCAGGCAGG CCAGGCACAC	120
GCATCCGCAG GACCCGCCGC ACCCGCCACG CAGACACGGA CGAGCCGCCG CGGTCAAGAT	180
GTTCAACCGC CGCGGTCAAG ATGTATGTGC CACCGACCCT CGCCCCGCTG GACGGACGGA	240
CGGACGCACG CACGCCGTCA GCGTCCACCG GTCACTGCCG CCGCCCACAG TGATGTCACC	300
CACGAAAGCA CACACGTAGA AGCGGACGCC GTGGTCAAGA TGTCTCTGCC ATCCCCACAG	360
GACGGACGGA CGGACTCCAC AAGGT	385